

Application No. 09/421,416
Amd. dated July 2, 2004
Reply to May 5, 2004 Notice of Allowance

Xerox Docket No. D/99198 - 311297

16. (Previously Presented) The method of claim 15, wherein the step of converting the second feature comprises the sub-steps of:

identifying each other digital document linked to by all digital documents within the collection of digital documents; and

creating the second vector having a number of dimensions equal to the number of other digital documents linked to by digital documents in the collection of digital documents, and the second vector further having as each element a numeric value representative of the number of links in the first digital document linking to each corresponding other digital document.

17. (Previously Presented) The method of claim 16, wherein the numeric value representative of the number of links in the first digital document linking to each corresponding other digital document is calculated as the token frequency weight of the corresponding link multiplied by the inverse context frequency weight of the corresponding link.

18. (Previously Presented) The method of claim 15, wherein the step of converting the second feature comprises the sub-steps of:

identifying each unique word associated with URLs defining hyperlinks in each digital document in the collection of digital documents;

counting the occurrences of each unique word in the URLs defining hyperlinks in the first digital document; and

creating the second vector having a number of dimensions equal to the number of unique words associated with the URLs defining hyperlinks in each digital document, and the second vector further having